Development and validation of a method to estimate body weight in critically ill children using length and mid-arm circumference measurements: The PAWPER XL-MAC system

M Wells, MB BCh, MSc (Med) (Emergency Medicine), FCEM (SA); **L N Goldstein**, MB BCh, MMed (Emergency Medicine), FCEM (SA), Cert Critical Care (SA); **A Bentley**, MB BCh, PhD

Division of Emergency Medicine, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

***Corresponding author:** M Wells (mike.wells@emergencymedicine.co.za)*

Background. Erroneous weight estimation during the management of emergency presentations in children may contribute to patient harm and poor outcomes. The PAWPER (Paediatric Advanced Weight Prediction in the Emergency Room) XL tape is an accurate length-based, habitus-modified weight estimation device, but is vulnerable to errors if subjective visual assessments of children's body habitus are incorrect or erratic.

Objective. Mid-arm circumference (MAC) has previously been used as a surrogate indicator of habitus, and the objective of this study was to determine whether MAC cut-off values could be used to predict habitus scores (HSs) to create an objective and standardised weight estimation methodology, the PAWPER XL-MAC method.

Methods. The PAWPER XL-MAC model was developed by creating MAC ranges for each HS in each weight segment of the tape. This model was validated against two samples, the National Health and Nutrition Examination Survey datasets and data from two previous PAWPER tape studies. The primary outcome measure was to achieve >70% of estimations within 10% of measured weight (PW10 >70%) and >95% within 20% of measured weight (PW20 >95%) for children aged 0 - 18 years.

Results. The PAWPER XL-MAC model achieved very high accuracy in the three validation datasets (PW10 79.2%, 79.0% and 81.9%) and a very low critical error rate (PW20 98.5%, 96.0% and 98.0%). This accuracy was maintained across all ages and in all habitus types, except for the severely obese.

Conclusions. The PAWPER XL-MAC model proved to be a very accurate, fully objective, standardised system in this study. It has the potential to be accurate across a wide variety of populations, even when used by those not experienced in visual assessment of habitus.

S Afr Med J 2017;107(11):1015-1021. DOI:10.7196/SAMJ.2017.v107i11.12505

It is important to predict the weight of children who require resuscitation, or any form of emergency medical care, accurately in order to provide appropriate doses of potentially life-saving medications.[1] Estimated weights have been shown to be extremely inaccurate in many settings, however, with some studies in underweight and obese populations having shown weight estimations to be inaccurate in up to 85% of children.[2-5] These are dangerous results that would inevitably lead to many critical medication errors. Even in populations without extreme prevalences of underweight or obesity, many commonly used methods, such as age-based formulae, are frequently inaccurate in more than half the children in whom they are used.[6] To continue to use such methods, which are known to be inaccurate, cannot be considered good medical practice.[7]

A variety of systems have been developed to improve the accuracy of weight prediction, but all have their limitations. The most accurate current weight estimation systems are the dual length- and habitus-based systems, such as the Mercy method and the PAWPER (Paediatric Advanced Weight Prediction in the Emergency Room) tape.[8,9] For these systems that rely on length to predict weight, it is important that habitus is quantified accurately in order to obtain a correct weight estimation in all children, including those who are underweight or obese.

The PAWPER XL tape (the next generation of the original PAWPER tape) was specifically designed to be used in emergencies and makes use of measurement of length as well as assessment of body habitus, either visually or with the assistance of figural reference images, to generate a rapid, accurate, calculation-free estimation of weight.[10,11] However, recent studies on the original PAWPER tape in very obese populations, and with novice users, failed to show the degree of accuracy reached in studies in less obese populations.[8,10,12-14] An important contributor to the reduced accuracy of the PAWPER tape was inaccurate assessment of body habitus in these obese children.[12,15]

The development of a less subjective and more standardised method of assessing habitus could therefore potentially improve the functioning of the PAWPER system in different populations and with users of different experience. Maximum objectivity can best be achieved with the use of a simple, easy-to-perform anthropometric measurement such as mid-arm circumference (MAC), since MAC has been shown to have a strong association with body weight, as well as body habitus, in children and adolescents.[9,16-18]

An additional issue is that, depending on the drugs to be used, there are situations where estimations of both total body weight (TBW) and ideal body weight (IBW) may be required to optimise drug administration. Obese children, in particular, could be at risk of toxicity or reduced therapeutic effect if drug doses are not corrected for body composition. In general, obese children should have lipophilic drugs (e.g. amiodarone) scaled to TBW and hydrophilic drugs scaled to IBW, even during emergency care.[19] A system that could provide simultaneous, accurate estimations of both TBW and IBW would be valuable.

Objective

This study was an attempt to use MAC as a rapid, objective tool to predict habitus and thus, in association with measured length, predict TBW (and IBW): the PAWPER XL-MAC system.

Methods

The steps followed in developing, calibrating and validating the model are shown in Fig. 1, along with an image of the tape developed to make these data practical during emergency care.

Development of the preliminary PAWPER XL-MAC model

The original PAWPER XL tape makes use of a visual assessment of habitus to assign a habitus score (HS) (HS1 - HS7, with HS1 representing an underweight child, HS3 an average child and HS7 a severely obese child), with the predicted weights associated with each HS based on weight-for-length growth chart centiles (see Fig. 1 for details). To make use of measurements of MAC to predict HS (and TBW), rather than a visual assessment of habitus, measurement ranges of MAC measurements needed to be created for each of the seven HSs in each segment of the tape. To do this, MAC-for-length centiles were created from growth chart data downloaded from the Centers for Disease Control (CDC) website.[20] These centiles were then matched with the weight-for-length centiles used to define each HS on the original PAWPER XL tape. Finally, the centiles were used to create the MAC measurement ranges used to define the HSs.

Once this theoretical preliminary model had been constructed, it was evaluated for accuracy and the MAC measurement ranges were fine-tuned where needed.

Calibration and refinement of the PAWPER XL-MAC model

The National Health and Nutrition Examination Survey (NHANES) datasets A - G (seven datasets from the 1999 - 2000 to 2011 - 2012 surveys) were downloaded from the CDC website.[21] The demographic and anthropometric data for all children ≤18 years of age were extracted. The specific variables retained included age, height or recumbent length, TBW, body mass index (BMI) and MAC. Cases with missing or incomplete data were excluded. The 2007 - 2008 and 2009 - 2010 surveys (datasets E and F) were pooled to be used in the first round of calibration. The other datasets were used for the final validation of the refined model. BMI-for-age z-scores were calculated using World Health Organization reference data for children aged <24 months and CDC reference data for children from 2 to 18 years of age.

Weight estimates were generated by the preliminary model using the measurements of length and MAC from the downloaded datasets. These weight estimates were then compared against the actual measured weight for each child. The accuracy of the estimations was evaluated in the entire sample as well as in each of the 34 weight segments of the PAWPER XL-MAC tape model.

The cut-off MAC values were adjusted (calibrated) until an acceptable degree of accuracy was obtained in each segment (see below for acceptable outcome measures). The outcome data before and after calibration are shown in Table I. The final model, the details of which are shown in Table II, was then subjected to a validation assessment.

Validation of the final PAWPER XL-MAC model

The final model was validated in three samples: the pooled unused NHANES data (datasets A - D and G) and data from two previous

Step 1: Development of model

MAC-for-length centiles were created by combining MAC-for-age and length-for-age growth chart data. These MAC-for-length centiles were matched with the weight-for-length centiles currently used by the PAWPER XL tape to define the seven HSs (see on the right). Finally, MAC ranges were derived to predict the habitus scores (and total body weight) in each segment of the tape.

HS1 5th centile
HS2 15th centile
HS3 50th centile
HS4 85th centile
HS5 95th centile
HS6 97th centile
HS7 99th centile

Step 2: Calibration of model

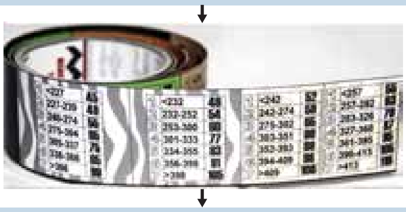
The preliminary model was then tested on the NHANES datasets E and F (from the 2007 - 2008 and 2009 - 2010 surveys). The MAC ranges were adjusted in each tape segment until a satisfactory degree of weight estimation accuracy was achieved.

Step 3: Validation of final model

The final model was tested in three independent datasets that had not been used in the model design. The first dataset was created by pooling data from the NHANES datasets A, B, C, D and G (the 1999 - 2000, 2001 - 2002, 2003 - 2004, 2005 - 2006 and 2011 - 2012 surveys). The other datasets were obtained from two previous weight estimation studies conducted in SA populations.

Step 4: Creation of the PAWPER XL-MAC tape

The validated data were used to compose the PAWPER XL-MAC tape. The tape (shown below) is 1 800 mm long with 13 colour zones and 34 IBW length segments ranging from 3 kg up to 70 kg. The beginning of the tape contains a printed scale to measure MAC. The zebra-stripe and giraffe-spot colour zones (with 55 kg to 70 kg segments) are shown in the image. The MAC ranges (in mm) that define the seven habitus scores and the corresponding predicted weights (in kg) can be seen. The IBW is predicted by the HS3 weight.



How to use the PAWPER XL-MAC tape

Step 1 Measure recumbent length with the tape.
Step 2 Measure the MAC and read the habitus score and estimated TBW directly off the tape at the measured length. Use TBW for dose scaling in all non-obese children.
Step 3 If required, read IBW directly off the tape (the HS3 weight). Use IBW only for obese children (HS5, HS6 and HS 7) and only for hydrophilic drug dose calculations.

Fig. 1. The methodology followed in the study. The process of creating the initial model, calibrating this model and then validating the final model in three independent samples is illustrated. The figure also shows how the validated data were converted into the full-length PAWPER XL-MAC tape, which could be used in clinical practice. Both TBW (adjusted for MAC-determined habitus) and IBW can be read directly off the tape. (MAC = mid-arm circumference; HS = habitus score; NHANES = National Health and Nutrition Examination Survey; SA = South African; IBW = ideal body weight; TBW = total body weight.)

PAWPER tape weight estimation studies in which MAC data were captured.[10,11] These two studies were prospective, cross-sectional studies conducted in two hospitals in Johannesburg, South Africa (SA). One of the hospitals serves a community of middle to upper socioeconomic status and the other a community of mostly low socioeconomic status. Study 1 enrolled 332 children from July 2014 to December 2014 and study 2 enrolled 300 children from August 2014 to January 2015. As in the calibration step, weight estimates were generated by the final model using the measurements of length and MAC from the validation datasets. These weight estimates were then compared against the measured weight for each child in the entire sample as well as in each of the individual segments.

Statistical analysis

The performance of the PAWPER XL-MAC model in predicting measured weight was evaluated using three major statistical measures: the mean percentage error (MPE) represented the

Table 1. Accuracy, bias and precision of the PAWPER XL-MAC model during development, before and after calibration and validation*

Dataset	*N*	PW10	PW20	MPE (LLOA, ULOA)
Accuracy of primary derivation model (before calibration)				
NHANES datasets E, F	4 664	53.2	86.2	−9.6 (−28.7, 9.6)
Accuracy of model after calibration				
NHANES datasets E, F	4 664	84.9	98.7	−2.1 (−16.8, 12.6)
Accuracy of final model in validation samples				
NHANES datasets A, B, C, D, G	13 134	81.9	98.0	−0.9 (−17.5, 15.7)
Study 1	332	79.2	98.5	1.5 (−14.5, 17.5)
Study 2	300	79.0	96.0	5.0 (−11.0, 21.0)
Original PAWPER tape methodology performance in SA validation samples				
Study 1	332	83.4	98.5	1.1 (−13.9, 16.1)
Study 2	300	88.0	97.7	1.4 (−12.5, 15.3)

MAC = mid-arm circumference; PW10 and PW20 = percentage of weight estimations falling within 10% and 20% of actual weight; MPE = mean percentage error; LLOA and ULOA = lower and upper limits of agreement; NHANES = National Health and Nutrition Examination Survey; SA = South Africa/n.
*The performance of the primary model before calibration is shown in the top section. PW10 and PW20 provide a quantification of accuracy. The MPE indicates the bias of the model, with a negative value denoting a bias to underestimation of weight. LLOA and ULOA provide an indication of precision: the ideal model should have 95% of estimations within ±20% of actual weight. The performance of the adjusted model (after calibration) in the derivation sample (NHANES datasets E and F from the 2007 - 2008 and 2009 - 2010 surveys) is shown in the second section. The performance of the final model in three new validation samples is shown in the bottom two sections. The NHANES dataset was derived from children in the USA (NHANES datasets A, B, C, D and G from the 1999 - 2000, 2001 - 2002, 2003 - 2004, 2005 - 2006 and 2011 - 2012 surveys), and the study 1 and study 2 datasets were derived from children in Johannesburg, SA. The final model achieved the acceptable outcome criteria of PW10 >70% and PW20 >95% in all three of the validation datasets. The validation of the final model was also performed independently in each of the 34 length segments of the tape, with the acceptable outcome criteria being achieved in every individual segment. The bottom panel shows the performance data of the original PAWPER methodology from the SA validation datasets, for comparison with the new model.

estimation bias; the 95% limits of agreement of the MPE quantified the estimation precision; and the percentage of weight estimations that fell within 10% (PW10) and 20% (PW20) of measured weight denoted overall accuracy. Subgroup analyses were performed on the pooled data for children of different weight status (underweight to severely obese).

All data were analysed using Microsoft Excel version 15.38, 2016 (Microsoft, USA) and Stata Statistical Software release 14, 2015 (StataCorp, USA).

Outcome measures

The primary outcome measure was the accuracy of the PAWPER XL-MAC model's estimation of weight when compared with measured weight. The secondary outcome measure was comparing the performance of the model with the accuracy of the original PAWPER tape method in the SA validation samples.

No previous weight estimation study has recommended an appropriate benchmark by which to characterise acceptable accuracy for a weight estimation system. It is, however, statistically important to have *a priori* criteria by which to judge the functioning of any methodology. We used a PW10 >70% together with a PW20 >95% to define acceptable accuracy of weight estimation. This was modified from the criteria proposed in an Australian biostatistical research report on the accuracy of the Broselow tape.[22] This target also matches the level of accuracy generally achieved by the most accurate existing weight estimation systems. A weight estimation error of >20% was considered critical because of the resultant high risk of medication error.

Results

Characteristics of study participants

The demographic characteristics of the children included in the calibration and validation studies are shown in Table III. The SA samples included a greater proportion of younger and underweight children than the NHANES dataset, although the study 1 sample contained a wide variety of body types with a high prevalence of both underweight and obese children. The study 2 sample was from an impoverished community with a very high prevalence of underweight children. Together the validation datasets provided children with a broad range of ages and body types in which to rigorously evaluate the model.

Validation of the PAWPER XL-MAC method

Primary outcome measures – achieving acceptable accuracy

As can be seen in Table I, the PAWPER XL-MAC method exceeded the acceptable outcome criteria (PW10 >70% and PW20 >95%) in each of the three validation samples overall, as well as in every segment-by-segment analysis. The best performance was in the 17 kg segment, with a PW10 of 95.7% and a PW20 of 99.5%, and the poorest performance was in the 60 kg segment, with a PW10 of 73.3% and a PW20 of 97.1%.

Bland-Altman plots illustrating the performance of the final model in the pooled validation samples, as well as for subgroups representing extremes of habitus, are shown in Fig. 2.

Subgroup analysis by extremes of body habitus (Table IV) showed that the primary outcome measures were fully met in five of six categories. Only in severely obese children was the accuracy poor (PW10 43.2%) and the critical error rate higher than the primary outcome measures permitted (PW20 80.6%).

Secondary outcome measures – comparison with original PAWPER methodology

The performance of the PAWPER XL-MAC method in the study 1 population, which included extremes of body habitus, was similar to the original results with the regular PAWPER XL methodology (see bottom of Table I). In the study 2 population, the PAWPER XL-MAC method performed acceptably well, but with a higher bias to overestimation of weight than with visual habitus assessment. The model was not as accurate as the original method (PW10 comparisons, odds ratio 2.2 (confidence interval 1.4 - 3.4); p=0.0012 (Fisher's exact test)). There was no difference in the critical error rate (PW20), however.

Discussion

Main findings

The PAWPER XL-MAC model satisfied the primary outcome measures by achieving a PW10 in >70% of children, and a critical error rate of <5% (PW20 >95%). The findings were consistent across all lengths and in all habitus types except for severely obese children, who accounted for just over 4% of the validation sample. The model was also accurate in two smaller SA validation samples, achieving a reasonably similar accuracy to the conventional PAWPER methodology. The diverse demographic and anthropometric charac-

Table 2. The detailed data describing the full PAWPER XL-MAC tape model*

		HS1		HS2			HS3			HS4			HS5			HS6			HS7	
Recumbent length (cm)	IBW (kg)	TBW (kg)	MAC upper limit (cm)	TBW (kg)	MAC lower limit (cm)	MAC upper limit (cm)	TBW (kg)	MAC lower limit (cm)	MAC upper limit (cm)	TBW (kg)	MAC lower limit (cm)	MAC upper limit (cm)	TBW (kg)	MAC lower limit (cm)	MAC upper limit (cm)	TBW (kg)	MAC lower limit (cm)	MAC upper limit (cm)	TBW (kg)	MAC lower limit (cm)
43.0 - 49.3	3	2.5	<9.4	2.8	9.4	10.3	3	10.4	11.5	3.2	11.6	12.7	3.5	12.8	13.9	4	14.0	15.1	4.5	15.2
49.4 - 54.8	4	3.5	<9.6	3.8	9.6	10.4	4	10.5	11.7	4.5	11.8	12.9	5	13.0	14.1	6	14.2	15.4	7	15.5
54.9 - 59.2	5	4	<11.0	4.5	11.0	11.4	5	11.5	12.4	5.5	12.5	13.4	6	13.5	15.4	7	15.5	16.4	8	16.5
59.3 - 62.9	6	5	<11.5	5.5	11.5	11.9	6	12.0	13.1	6.5	13.2	14.2	7	14.3	15.2	8	15.3	16.4	9	16.5
63.0 - 66.4	7	6	<11.5	6.5	11.5	12.9	7	13.0	14.4	7.5	14.5	15.3	8	15.4	16.4	9	16.5	17.4	10	17.5
66.5 - 70.1	8	7	<12.5	7.5	12.5	13.6	8	13.7	14.4	8.5	14.5	15.4	9.5	15.5	16.9	10	17.0	18.4	11	18.5
70.2 - 73.9	9	7.5	<12.6	8.5	12.6	14.1	9	14.2	15.4	10	15.5	16.2	11	16.3	18.1	12	18.2	19.9	13	20.0
74.0 - 78.2	10	8.5	<12.7	9.5	12.7	14.4	10	14.5	15.4	11	15.5	17.0	12	17.1	18.6	13	18.7	20.4	14	20.5
78.3 - 83.1	11	9	<12.7	10	12.7	14.6	11	14.7	15.4	12	15.5	17.0	13	17.1	18.9	14	19.0	20.4	15	20.5
83.2 - 88.0	12	10	<13.0	11	13.0	14.8	12	14.9	15.7	13	15.8	17.1	14	17.2	18.9	16	19.0	20.4	18	20.5
88.1 - 92.6	13	11	<13.5	12	13.5	14.8	13	14.9	15.8	14	15.9	17.2	15	17.3	19.0	17	19.1	20.9	19	21.0
92.7 - 96.8	14	12	<14.2	13	14.2	15.1	14	15.2	16.5	15	16.6	17.6	17	17.7	19.3	19	19.4	20.9	21	21.0
96.9 - 100.6	15	13	<14.4	14	14.4	15.6	15	15.7	16.5	16	16.6	17.6	18	17.7	19.4	20	19.5	21.4	22	21.5
100.7 - 103.9	16	14	<14.8	15	14.8	15.7	16	15.8	17.0	17	17.1	17.9	19	18.0	19.7	21	19.8	22.4	23	22.5
104.0 - 107.0	17	15	<15.3	16	15.3	15.9	17	16.0	17.4	18	17.5	18.0	20	18.1	19.9	22	20.0	22.4	24	22.5
107.1 - 110.0	18	16	<15.6	17	15.6	16.6	18	16.7	17.8	19	17.9	18.6	21	18.7	21.9	24	22.0	23.4	26	23.5
110.1 - 113.2	19	16	<15.6	18	15.6	16.6	19	16.7	17.8	21	17.9	20.2	23	20.3	22.5	25	22.6	24.4	28	24.5
113.3 - 116.5	20	17	<15.6	19	15.6	17.5	20	17.6	18.4	22	18.5	20.2	24	20.3	22.9	26	23.0	24.4	29	24.5
116.6 - 120.6	22	20	<16.3	21	16.3	17.7	22	17.8	19.1	24	19.2	21.2	27	21.3	23.1	29	23.2	24.8	32	24.9
120.7 - 125.4	24	21	<16.5	23	16.5	18.2	24	18.3	19.4	26	19.5	21.8	29	21.9	23.4	32	23.5	24.8	35	24.9
125.5 - 129.6	26	23	<17.5	25	17.5	18.5	26	18.6	20.4	28	20.5	22.0	31	22.1	24.4	34	24.5	25.4	37	25.5
129.7 - 133.3	28	24	<17.8	26	17.8	18.7	28	18.8	20.7	30	20.8	22.3	33	22.4	25.6	36	25.7	26.6	40	26.7
133.4 - 136.6	30	26	<17.9	28	17.9	19.4	30	19.5	22.0	33	22.1	22.9	36	23.0	25.6	40	25.7	26.6	44	26.7
136.7 - 139.8	32	28	<19.2	30	19.2	20.4	32	20.5	22.3	35	22.4	23.9	39	24.0	25.6	42	25.7	26.9	47	27.0
139.9 - 143.2	34	29	<19.2	32	19.2	20.7	34	20.8	23.2	38	23.3	24.6	42	24.7	26.5	46	26.6	28.2	50	28.3
143.3 - 146.5	36	29	<19.5	33	19.5	21.0	36	21.1	23.3	39	23.4	24.6	43	24.7	26.9	48	27.0	29.9	52	30.0
146.6 - 149.8	38	33	<20.2	36	20.2	21.6	38	21.7	23.3	41	23.4	24.9	45	25.0	26.9	50	27.0	28.6	58	28.7
149.9 - 153.1	40	36	<20.8	39	20.8	22.5	40	22.6	24.1	48	24.2	26.1	52	26.2	27.8	57	27.9	29.9	65	30.0
153.2 - 158.0	45	38	<21.1	43	21.1	23.4	45	23.5	25.5	55	25.6	28.1	64	28.2	31.5	72	31.6	35.1	84	35.2
158.1 - 165.0	50	40	<22.0	45	22.0	23.8	50	23.9	26.1	60	26.2	30.1	72	30.2	32.7	79	32.8	35.3	87	35.4
165.1 - 170.0	55	45	<22.7	49	22.7	23.9	55	24.0	27.4	65	27.5	30.4	75	30.5	33.7	85	33.8	36.6	99	36.7
170.1 - 174.0	60	46	<23.2	54	23.2	25.2	60	25.3	30.0	77	30.1	33.3	83	33.4	35.5	91	35.6	39.8	105	39.9
174.1 - 177.0	65	52	<24.2	59	24.2	27.4	65	27.5	30.2	80	30.3	35.1	89	35.2	39.3	98	39.4	40.9	108	41.0
177.1 - 180.0	70	55	<25.7	63	25.7	28.2	70	28.3	32.6	87	32.7	36.0	95	36.1	39.5	105	39.6	41.3	116	41.4

MAC = mid-arm circumference; HS = habitus score (HS1 - HS7, HS1 representing an underweight child, HS3 an average child and HS7 a severely obese child); IBW = ideal body weight; TBW = total body weight.
*The child's recumbent length determines into which length segment they would fall. The HS3 weight for each segment reflects the IBW for a child of that length (i.e. a child with a measured recumbent length of between 43.0 and 49.3 cm would be predicted to have an IBW of 3 kg). The upper and lower limits of MAC measurements that were used to define each HS within each length segment are shown. The final adjusted or predicted TBW for children of below or above average weight-for-length are shown for each HS in each segment (i.e. a child with a recumbent length of 137.0 cm and an MAC of 26.0 cm would be predicted to have an IBW of 32 kg but a TBW of 42 kg – an obese child). A full-length tape was designed from these data so that both TBW and IBW could be rapidly determined during emergency care.

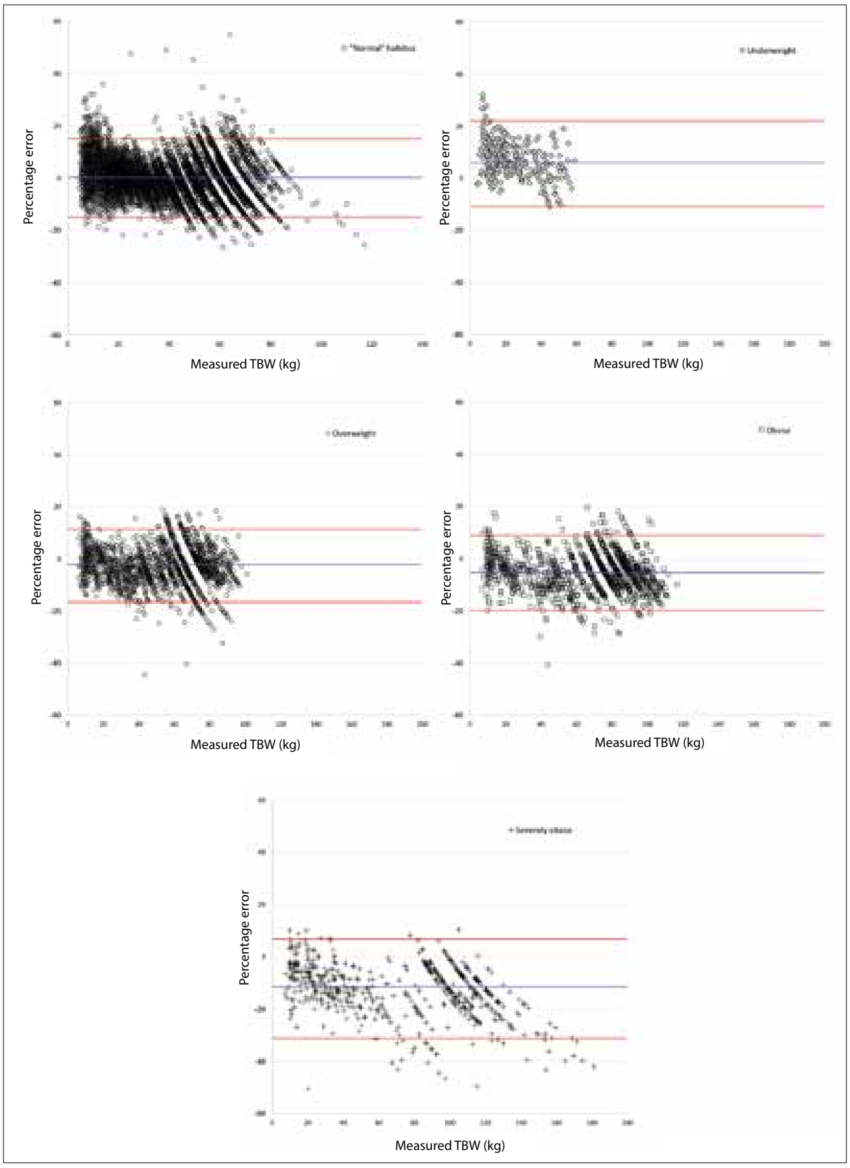


Fig. 2. Bland-Altman plots of percentage error against actual weight for the pooled validation datasets for children with 'average' and extremes of habitus. The blue lines represent the mean percentage error (trueness or bias) and the red lines represent the 95% limits of agreement (i.e. the range in which 95% of the estimates fall). (TBW = total body weight.)

teristics of the validation samples suggest that this methodology may be equally accurate in other similar populations.

The strength of the model derived in this study was its consistency and accuracy in >95% of the sample. The progression in expectations of weight estimation systems from those that achieve accuracy in <25% of children to the newer methodologies in which >70% receive an accurate estimation and <5% have a critical error, is encouraging.[9,23] It is possible that severely obese children will remain a challenge with regard to accurate weight estimation and appropriate drug dose determination, however, especially children with BMI-for-age *z*-scores well above 5.

Visual assessment of HS

Although the PAWPER tape has been shown to outperform other methods of weight estimation in children in some studies, erratic subjective assessment of body habitus by users has resulted in less-than-desirable accuracy in other studies and populations.[12-14] Accurate and repeatable anthroposcopic assessment of habitus is possible by experienced practitioners, but assessment has been shown to be less reliable when performed by novices.[24,25] Furthermore, healthcare professionals, and parents, have more difficulty in recognising overweight and obese children than underweight children, which would lead to an underestimation of HS when using the regular PAWPER XL tape methodology.[15,26] Figural reference images, validated against BMI, have been used successfully to reduce the subjectivity of anthroposcopic assessment of habitus, mostly in body image research.[27] Research on the use of figural reference images with the PAWPER XL tape has shown good results, with equal accuracy when compared with experts in visual gestalt assessment.[10] Although the use of figural reference images has been shown to improve visual assessment of habitus, it remains somewhat subjective.[10,28] A fully objective method could therefore have an important role, even if used as a secondary or confirmatory technique for the PAWPER XL tape.

In this study, the accuracy achieved by the PAWPER XL-MAC model was comparable to that obtained by visual assessment of habitus. Although the accuracy of the original PAWPER tape using visual assessment of habitus was statistically significantly better than the new PAWPER XL-MAC model in the study 2 population, the new model still performed very well. The evidence suggests that extraordinarily high levels of accuracy are achievable in underweight and normal-weight children using visual assessment, especially with skilled users, but that this is more difficult in obese children.[15,29] This disparity makes gestalt visual assessment vulnerable to error and may potentially increase cognitive stress, especially in inexperienced users. An objective, standardised, easy-to-use method is of value.

MAC and body habitus

MAC has a strong evidence base supporting its value as a surrogate for body habitus and is a reliable measurement with high inter-user agreement.[30] When used as a single variable, MAC is more accurate than age in predicting weight.[31] It is the habitus-surrogate measurement used with the Mercy method, which has been shown to be very accurate in a wide range of underweight and obese populations.[9,32] The accuracy of the PAWPER XL-MAC model in this study was similar to, or better than, the Mercy method outcomes reported previously (PW10 63.9 - 80.1% range), but far better than MAC used as a single variable (PW10 15.5 - 44.2%).[9,32] The use of MAC in the PAWPER XL-MAC model to predict HS was accurate and objective, fulfilling the main objectives of this study. This model does not change how IBW may be rapidly estimated in obese children (the HS3 weight is used, which is printed on the tape), if required. With weight estimation systems designed for use during emergency care, however, how the systems work may be as important as how accurate they are.[33]

Table 3. Demographic characteristics of children in the derivation and validation datasets*

	Calibration dataset, NHANES: USA pooled data (datasets E, F)	Validation datasets		
		NHANES: USA pooled data, datasets A, B, C, D, G	Study 1: SA, low- and middle-income groups	Study 2: SA, low-income group
N	4 664	13 134	332	300
Age (yr), median (IQR)	10.0 (4.0 - 14.0)	10.0 (4.0 - 14.0)	7.2 (4.5 - 9.3)	4.3 (2.3 - 7.0)
Sex male, *n* (%)	2 326 (49.9)	6 483 (49.4)	154 (46.7)	152 (50.7)
Length (cm), median (IQR)	145.2 (104.0 - 161.3)	142.4 (104.3 - 162)	122.0 (106.0 - 137.0)	104.0 (87.0 - 122.0)
Weight (kg), median (IQR)	40.5 (17.2 - 59.3)	38.1 (17.3 - 58.6)	23.4 (17.6 - 33.8)	16.2 (12.3 - 22.2)
BMI (kg/m^2), median (IQR)	18.7 (16.4 - 22.8)	18.7 (16.5 - 22.5)	16.7 (15.2 - 18.8)	15.2 (14.2 - 16.4)
z-score, median (IQR)	0.6 (–0.2 - 1.5)	0.5 (–0.3 - 1.4)	0.4 (–0.5 - 1.1)	–0.7 (–1.7 - 0.2)
Underweight, *n* (%)	77 (2.0)	348 (2.6)	23 (7.0)	43 (18.3)
Overweight, *n* (%)	522 (13.6)	1 686 (12.8)	38 (11.5)	12 (5.1)
Obese, *n* (%)	368 (9.6)	1 074 (8.2)	20 (6.1)	4 (1.7)
Severely obese, *n* (%)	161 (4.2)	572 (4.4)	15 (4.5)	1 (0.4)

NHANES = National Health and Nutrition Examination Survey; SA = South Africa; BMI = body mass index.
*There were substantial differences between the ages and weight status of the three validation samples. Study 2, which comprised mainly children from a low-income group, had a much higher number of underweight children than the other datasets. Study 1, which included children from both low- and middle-income groups, had a greater variety of body types than the other two datasets. The three datasets provided a reasonable variation of age groups and habitus types in which to assess the model performance.

Table 4. Subgroup analysis of the PAWPER XL-MAC model in the pooled validation datasets, by body composition*

Description	BMI-for-age *z*-score	*n* (%)	PW10	PW20
Underweight	≤2	414 (3.0)	71.3	96.3
Thin	–1.4 - –1.9	511 (3.7)	82.2	98.4
'Average'	–1.3 - 1.3	9 419 (68.4)	85.7	99.1
Overweight	1.4 - 1.9	1 736 (12.6)	82.7	98.6
Obese	2 - 2.4	1 098 (8.0)	73.1	98.1
Severely obese	≥2.5	588 (4.3)	43.2	80.6

BMI = body mass index; PW10 and PW20 = percentage of weight estimations falling within 10% and 20% of actual weight.
*There were 13 766 total available *z*-scores for inclusion. The model failed to achieve acceptable accuracy only in the severely obese group of children – the *z*-scores in this group ranged from 2.5 to 7.9, with the children's weight ranging up to 181 kg.

Constraints on the use of weight estimation systems during resuscitation

The cognitive stresses incurred by healthcare providers during emergency care reduce their ability to perform even the simplest mental tasks.[33,34] Calculation errors are particularly common, which may make any system that requires arithmetic vulnerable when used during emergencies.[35] The PAWPER XL-MAC model was therefore developed so that MAC cut-off values (ranges) could be read rapidly off the tape to determine the HS, without requiring calculations – a cognitively neutral process. This is a more complex process than visual assessment of habitus, but also one that would not instil doubts about the accuracy of habitus assessment. Adequate training would clearly be required to use the PAWPER XL-MAC system accurately – even the most basic of weight estimation techniques (such as the Broselow tape) have been shown to be performed very poorly when used by individuals with no training.[36,37] Given the somewhat increased complexity of the PAWPER XL-MAC system, a staged or protocol-driven procedure for weight estimation and weight estimation training may be required.

Staged approach to weight estimation

The use of the PAWPER XL-MAC method would be most suitable if implemented in a staged approach to weight estimation, timing the procedures to the priorities and dictates of the emergency medical care. A visual or image-assisted PAWPER XL tape weight could be obtained immediately emergency, life-saving care is initiated, followed by a confirmatory measurement-based PAWPER XL-MAC estimation several minutes later once exact drug doses are required, prior to their administration. Weight estimation practices should also be incorporated into paediatric resuscitation courses and regular simulation training. This concept would need additional evaluation in simulation-based training.

TBW and IBW estimations

It has been speculated that the use of inaccurate or inappropriate weight estimations may worsen outcomes in obese children.[38] Since there is now a reasonable consensus on which drugs should be scaled to TBW and which to IBW in obese children, these should be applied during emergency care.[39] The PAWPER XL-MAC model could facilitate the estimation of both, in keeping with recent guidelines. How this may provide benefits for care still needs to be investigated.

Study limitations

There were relatively small numbers of infants weighing <6 kg in the calibration and validation datasets, so it is uncertain whether the accuracy of the PAWPER XL-MAC technique can be generalised to this weight group, and further confirmation will be required. Measurement of MAC in a supine child receiving emergency care by a stressed healthcare provider may not be as accurate as that performed on a seated child by an anthropometry expert. This also requires further research in real-world or simulation situations.

Conclusions

The consistent level of high accuracy achieved by the PAWPER XL-MAC method across a broad age range of children in this study exceeded the stringent outcome measures. Weight estimates were also accurate in children of all habitus types except for the severely obese (BMI-for-age z-score ≥2.5), in whom accurate weight estimation still proves challenging. Critical error rates were generally very low. The PAWPER XL-MAC method achieved similar accuracy of weight estimation to the original PAWPER methodology (using gestalt visual assessment) from previous studies.

The objectiveness of the PAWPER XL-MAC technique may make it more standardised and objective for novice users and those not experienced in or skilled at anthroposcopic habitus assessment. The immediate feedback with regard to HS assessment could also assist users in learning the visual method of habitus scoring.

The validation samples comprised a very diverse group of children, with the model performing well in almost all the subgroups, supporting speculation that the model may be accurate in other populations with a similar demographic composition. However, the PAWPER XL-MAC needs prospective testing in a variety of populations to establish its true accuracy and generalisability.

Acknowledgements. None.

Author contributions. The authors all contributed to the study design, analysis and interpretation of data and drafting and writing of the article.

Funding. None.

Conflicts of interest. None.

1. Kaufmann J, Laschat M, Wappler F. Medication errors in pediatric emergencies: A systematic analysis. Dtsch Arztebl Int 2012;109(38):609-616. https://doi.org/10.3238/arztebl.2012.0609
2. Clark MC, Lewis RJ, Fleischman RJ, Ogunniyi AA, Patel DS, Donaldson RI. Accuracy of the Broselow tape in South Sudan, 'the hungriest place on earth'. Acad Emerg Med 2016;23(1):21-28. https://doi.org/10.1111/acem.12854
3. Bowen L, Zyambo M, Snell D, Kinnear J, Bould MD. Evaluation of the accuracy of common weight estimation formulae in a Zambian paediatric surgical population. Anaesthesia 2017;72(4):470-478. https://doi.org/10.1111/anae.13780
4. Marlow R, Lo D, Walton L. Accurate paediatric weight estimation by age: Mission impossible? Arch Dis Child 2011;96(Suppl 1):A1-A2. https://doi.org/10.1136/adc.2011.212563.3
5. Cattermole GN, Leung MP, So HK, Mak PS, Graham CA, Rainer TH. Age-based formulae to estimate children's weight in the emergency department. Emerg Med J 2011;28(5):390-396. https://doi.org/10.1136/emj.2009.090357
6. Wells M, Goldstein L, Bentley A. It is time to abandon age-based emergency weight estimation in children! A failed validation of 20 different age-based formulas. Resuscitation 2017;116(7):73-83. https://doi.org/10.1016/j.resuscitation.2017.05.018
7. Luscombe M, Owens B. Weight estimation in resuscitation: Is the current formula still valid? Arch Dis Child 2007;92(5):412-415. https://doi.org/10.1136/adc.2006.107284
8. Wells M, Coovadia A, Kramer E, Goldstein L. The PAWPER tape: A new concept tape-based device that increases the accuracy of weight estimation in children through the inclusion of a modifier based on body habitus. Resuscitation 2013;84(2):227-232. https://doi.org/10.1016/j.resuscitation.2012.05.028
9. Abdel-Rahman SM, Ridge AL. An improved pediatric weight estimation strategy. Open Med Devices J 2012;4:87-97. https://doi.org/10.2174/1875181401204010087
10. Georgoulas V, Wells M. The PAWPER tape and the Mercy method outperform other methods of weight estimation in children in South Africa. S Afr Med J 2016;106(9):933-939. https://doi.org/10.7196/SAMJ.2016.v106i9.10572
11. Wells M, Goldstein L, Bentley A. A validation study of the PAWPER XL tape: Accurate estimation of both total and ideal body weight in children up to 16 years of age. Trauma Emerg Care 2017;2(4):1-8. https://www.oatext.com/pdf/TEC-2-141.pdf (accessed 27 September 2017).
12. Chavez H, Peterson R, Lo K, Arel M. Weight estimation in an inner city pediatric emergency department: The effect of obesity. Am J Emerg Med 2015;33(10):1364-1367. https://doi.org/10.1016/j.ajem.2015.07.040
13. Garcia CM, Meltzer JA, Chan KN, Cunningham SJ. A validation study of the PAWPER (Pediatric Advanced Weight Prediction in the Emergency Room) tape – a new weight estimation tool. J Pediatr 2015;167(1):173-177.e1. https://doi.org/10.1016/j.jpeds.2015.04.006
14. O'Leary F, John-Denny B, McGarvey K, Hann A, Pegiazoglou I, Peat J. Estimating the weight of ethnically diverse children attending an Australian emergency department: A prospective, blinded, comparison of age-based and length-based tools including Mercy, PAWPER and Broselow. Arch Dis Child 2017;102(1):46-52. https://doi.org/10.1136/archdischild-2016-310917
15. Goldstein L, Wells M. Fat is the new normal. J Pediatr 2016;168:257. https://doi.org/10.1016/j.jpeds.2015.09.049
16. Cattermole GN, Leung M, Mak P, So HK, Graham CA, Rainer TH. Children's weights correlate more strongly with mid-arm circumference (MAC) than with age, height or foot-length. J Emerg Med 2009;37(2):228-229. https://doi.org/10.1016/j.jemermed.2009.06.051
17. Cattermole GN, Graham CA, Rainer TH. Mid-arm circumference can be used to estimate weight of adult and adolescent patients. Emerg Med J 2016;34(4). https://doi.org/10.1136/emermed-2015-205623
18. Whitfield KC, Wozniak R, Pradinuk M, et al. Anthropometric measures are simple and accurate paediatric weight-prediction proxies in resource-poor settings with a high HIV prevalence. Arch Dis Child 2017;102(1):10-16. https://doi.org/10.1136/archdischild-2015-309645
19. Kendrick JG, Carr RR, Ensom MH. Pediatric obesity: Pharmacokinetics and implications for drug dosing. Clin Ther 2015;37(9):1897-1923. https://doi.org/10.1016/j.clinthera.2015.05.495
20. National Center for Health Statistics, Centers for Disease Control and Prevention. Anthropometric Reference Data for Children and Adults: United States, 2007 - 2010. Washington, DC: CDC/NCHS, 2012.
21. National Center for Health Statistics, Centers for Disease Control and Prevention. National Health and Nutrition Examination Survey Data. 2016. https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/default.aspx (accessed 27 September 2017).
22. Stewart D. Accuracy of the Broselow tape for estimating paediatric weight in two Australian emergency departments. University of Sydney, 2009. https://ses.library.usyd.edu.au//bitstream/2123/6265/1/Declan Stewart WPP final.pdf (accessed 27 September 2017).
23. Seddon C, Lockitt L, Dhanjal S, Eisenhut M. Validation of advanced paediatric life support formulas for weight calculation in a multiethnic population. International Scholarly Research Notices. Pediatrics 2012(2012), article ID 869634. https://doi.org/10.5402/2012/869634
24. Lockner DW, Heyward VH, Griffin SE, Marques MB, Stolarczyk LM, Wagner DR. Cross-validation of modified fatness-specific bioelectrical impedance equations. Int J Sport Nutr 1999;9(1):48-59. https://doi.org/10.1123/ijsn.9.1.48
25. Black K, Barnett P, Wolfe R, Young S. Are methods used to estimate weight in children accurate? Emerg Med 2002;14(2):160-165. https://doi.org/10.1046/j.1442-2026.2002.00311.x
26. Warschburger P, Kroller K. Childhood overweight and obesity: Maternal perceptions of the time for engaging in child weight management. BMC Public Health 2012;12:295. https://doi.org/10.1186/1471-2458-12-295
27. Truby H, Paxton SJ. The Children's Body Image Scale: Reliability and use with international standards for body mass index. Br J Clin Psychol 2008;47(Pt 1):119-124.
28. Britnell S, Taylor S, Koziol-McLain J. Emergency weight estimation lookup tables for New Zealand children aged 5 - 10 years. Emerg Med Australas 2016;28(5):558-563. https://doi.org/10.1111/1742-6723.12637
29. Twarog JP, Politis MD, Woods EL, Daniel LM, Sonneville KR. Is obesity becoming the new normal? Age, gender and racial/ethnic differences in parental misperception of obesity as being 'about the right weight'. Int J Obes (Lond) 2016;40(7):1051-1055. https://doi.org/10.1038/ijo.2016.40
30. WHO Multicentre Growth Reference Study Group. Reliability of anthropometric measurements in the WHO Multicentre Growth Reference Study. Acta Paediatr 2006;450(Suppl):38-46. https://doi.org/10.1111/j.1651-2227.2006.tb02374.x
31. Cattermole GN, Leung PY, Mak PS, Graham CA, Rainer TH. Mid-arm circumference can be used to estimate children's weights. Resuscitation 2010;81(9):1105-1110. https://doi.org/10.1016/j.resuscitation.2010.05.015
32. Dicko A, Alhousseini ML, Sidibé B, Traoré M, Abdel-Rahman SM. Evaluation of the Mercy weight estimation method in Ouelessebougou, Mali. BMC Public Health 2014;14:270. https://doi.org/10.1186/1471-2458-14-270
33. Luten R, Wears RL, Broselow J, Croskerry P, Joseph MM, Frush K. Managing the unique size-related issues of pediatric resuscitation: Reducing cognitive load with resuscitation aids. Acad Emerg Med 2002;9(8):840-847. https://doi.org/10.1197/aemj.9.8.840
34. Lancaster L. Mission impossible – minimising the terror of paediatric resuscitation for staff in the ED. Australas Emerg Nurs J 2002;5(3):16-20. https://doi.org/10.1016/S1328-2743(02)80004-9
35. Pentin J, Smith J. Drug calculations: Are they safer with or without a calculator? Br J Nurs 2006;15(14):778-781. https://doi.org/10.12968/bjon.2006.15.14.21582
36. Abdel-Rahman A, Jacobsen R, Watts J, et al. Comparative performance of pediatric weight estimation techniques: A human factor errors analysis. Pediatr Emerg Care 2015;33(8):548-552. https://doi.org/10.1097/PEC.0000000000000543
37. Hohenhaus SM, Frush KS. Pediatric patient safety in resuscitation simulations: Lessons learned from studying clinician behavior. Pediatr Emerg Care 2005;21(1):71-75. https://doi.org/10.1097/01.pec.0000150992.49722.60
38. Srinivasan V, Nadkarni VM, Helfaer MA, Carey SM, Berg RA, American Heart Association National Registry of Cardiopulmonary Resuscitation I. Childhood obesity and survival after in-hospital pediatric cardiopulmonary resuscitation. Pediatrics 2010;125(3):e481-e488. https://doi.org/10.1542/peds.2009-1324
39. Ross EL, Heizer J, Mixon MA, et al. Development of recommendations for dosing of commonly prescribed medications in critically ill obese children. Am J Health Syst Pharm 2015;72(7):542-556. https://doi.org/10.2146/ajhp140280

Accepted 13 July 2017.